GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 23, 2001, 10:48:45 ; Search time 23.25 Seconds
(without alignments)
892.293 Million cell updates/sec

Title: US-09-195-368-1
Perfect score: 951
Sequence: 1 MCLSHLENMPLSHSRTQGAQ.....VLKNNTYWGIILLANPQFIS 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	951	100.0	177	4	Q9UNG2	Q9ung2 homo sapien
2	906	95.3	169	4	O95852	O95852 homo sapien
3	108.5	11.4	272	13	Q9I8D8	Q9i8d8 gallus gall
4	85.5	9.0	415	4	Q9UGY8	Q9ugy8 homo sapien
5	84	8.8	579	12	Q65605	Q65605 bovine immu
6	82	8.6	1188	5	O96143	O96143 plasmodium
7	79.5	8.4	665	5	Q17701	Q17701 caenorhabdi
8	79.5	8.4	851	10	Q9LFL1	Q9lfl1 arabidopsis
9	78	8.2	326	12	P89064	P89064 rotavirus.
10	78	8.2	564	2	P75041	P75041 mycoplasma
11	77.5	8.1	352	2	Q9KM00	Q9km00 vibrio chol
12	77.5	8.1	3614	5	Q9W4E2	Q9w4e2 drosophila
13	77	8.1	326	12	Q9QS10	Q9qs10 human rotav
14	77	8.1	433	5	Q26197	Q26197 plasmodium
15	77	8.1	748	3	Q9P3D8	Q9p3d8 neurospora
16	77	8.1	1424	5	Q9N3L5	Q9n3l5 caenorhabdi
17	76	8.0	326	12	Q98022	Q98022 human rotav
18	76	8.0	326	12	Q98025	Q98025 human rotav
19	76	8.0	326	12	Q9IP76	Q9ip76 human rotav
20	76	8.0	525	2	O51223	O51223 borrelia bu
21	76	8.0	606	12	P87521	P87521 bovine immu
22	75.5	7.9	222	10	Q40766	Q40766 picea abies
23	75.5	7.9	314	7	Q31417	Q31417 latimeria c
24	75.5	7.9	686	5	Q19467	Q19467 caenorhabdi
25	75	7.9	575	12	Q9YW23	Q9yw23 melanoplus
26	74.5	7.8	222	10	Q9ZTY6	Q9zty6 pinus resin
27	74.5	7.8	271	10	Q9ZT29	Q9zt29 nicotiana t
28	74.5	7.8	465	3	Q9Y7T6	Q9y7t6 schizosacch
29	74.5	7.8	471	2	Q06959	Q06959 vibrio chol
30	74.5	7.8	985	5	O01590	O01590 caenorhabdi
31	74	7.8	326	12	Q98032	Q98032 human rotav
32	74	7.8	326	12	Q9IP71	Q9ip71 human rotav
33	74	7.8	881	10	Q9LJS2	Q9ljs2 arabidopsis
34	74	7.8	1895	12	Q9QAY3	Q9qay3 yaba monkey
35	73.5	7.7	218	10	Q9SAR6	Q9sar6 picea maria
36	73.5	7.7	222	10	Q9ZRC6	Q9zrc6 picea maria
37	73.5	7.7	222	10	Q9S7I9	Q9s7i9 picea maria
38	73.5	7.7	345	2	Q9ZKC1	Q9zkc1 helicobacte
39	73.5	7.7	504	10	Q9M1N3	Q9m1n3 arabidopsis
40	73.5	7.7	550	3	Q9Y788	Q9y788 cercospora
41	73	7.7	222	5	Q9V4G2	Q9v4g2 drosophila
42	73	7.7	326	12	Q98047	Q98047 human rotav
43	73	7.7	389	2	Q9X4V0	Q9x4v0 streptococc
44	73	7.7	494	12	Q86280	Q86280 rhesus rota
45	72.5	7.6	160	10	Q9LI85	Q9li85 arabidopsis

ALIGNMENTS

RESULT 1
Q9UNG2
ID Q9UNG2 PRELIMINARY; PRT; 177 AA.
AC Q9UNG2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE GLUCOCORTICOID-INDUCED TNFR-RELATED PROTEIN LIGAND.
GN TNFSF18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL VEIN;
RA Gurney A.L., Marsters S.A., Huang A., Pitti R.M., Mark M.,
RA Baldwin D.T., Gray A.M., Dowd P., Brush J., Heldens S., Schow P.,
RA Goddard A.D., Wood W.I., Baker K.P., Godowski P.J., Ashkenazi A.;
RT "Identification of a new member of the tumor necrosis factor family
RT and its receptor, a human ortholog of mouse GITR.";
RL Curr. Biol. 0:0-0(1999).
DR EMBL; AF125303; AAD22634.1; -.
SQ SEQUENCE 177 AA; 20307 MW; 3D78CE6B90F4C9E3 CRC64;

Query Match 100.0%; Score 951; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 3.7e-84;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MCLSHLENMPLSHSRTQGAQRSSWKLWLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MCLSHLENMPLSHSRTQGAQRSSWKLWLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMA 60

Qy      61 KFGPLPSKWQMASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KFGPLPSKWQMASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKN 120

Qy     121 KDMIQTLTNKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANPQFIS 177
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 KDMIQTLTNKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANPQFIS 177
```

```
RESULT   2
O95852
ID   O95852      PRELIMINARY;      PRT;   169 AA.
AC   O95852;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   AITR LIGAND.
GN   TL6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99156876; PubMed=10037686;
RA   Kwon B., Yu K.Y., Ni J., Yu G.L., Jang I.K., Kim Y.J., Xing L.,
RA   Liu D., Wang S.X., Kwon B.S.;
RT   "Identification of a novel activation-inducible protein of the tumor
RT   necrosis factor receptor superfamily and its ligand.";
RL   J. Biol. Chem. 274:6056-6061(1999).
DR   EMBL; AF117713; AAD19695.1; -.
SQ   SEQUENCE   169 AA;  19379 MW;  D6AC467D94FE17C3 CRC64;


  Query Match             95.3%;  Score 906;  DB 4;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-80;
  Matches  169;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 MPLSHSRTQGAQRSSWKLWLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMAKFGPLPSK 68
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPLSHSRTQGAQRSSWKLWLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMAKFGPLPSK 60

Qy      69 WQMASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMIQTLT 128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 WQMASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMIQTLT 120

Qy     129 NKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANPQFIS 177
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANPQFIS 169


RESULT   3
Q9I8D8
ID   Q9I8D8      PRELIMINARY;      PRT;   272 AA.
AC   Q9I8D8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE CD154 (CD40 LIGAND).
GN   CD154.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN LINE 0; TISSUE=SPLEEN;
RA   Tregaskes C.A., Young J.R., Burnside J.;
RT   "cloning of a putative chicken CD40 ligand.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ243435; CAB95748.1; -.
SQ   SEQUENCE   272 AA;  30862 MW;  5409F24A8E53CCD7 CRC64;


  Query Match             11.4%;  Score 108.5;  DB 13;  Length 272;
  Best Local Similarity   29.8%;  Pred. No. 0.0091;
  Matches   36;  Conservative   22;  Mismatches   44;  Indels   19;  Gaps    5;

Qy      68 KWQMASSEPPCVNKVS--DWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLY------- 118
           || | :|  |  : :|  : ||:: : ||| || ||:       ||| : :|

Db     153 KW-MTTSYAPTSSLISYHEGKLKVEKAGLYYIYSQVSFCTKAAASAPFTLYIYLYLPMEE 211

Qy     119 -----KNKDMIQTLTNKSKIQNV--GGTYELHVGDT--IDLIFNSEHQVLKNNTYWGIIL 169
                | |   | |   ::|::  || :||  ||   :::  ::   |   |||:|:
Db     212 DRLLMKGLDTHSTSTALCELQSIREGGVFELRQGDMVFVNVTDSTAVNVNPGNTYFGMFK 271

Qy     170 L 170
           |
Db     272 L 272


RESULT   4
Q9UGY8
ID   Q9UGY8      PRELIMINARY;      PRT;   415 AA.
AC   Q9UGY8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   DJ355C18.1 (KIAA0027) (FRAGMENT).
GN   DJ355C18.1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cobley V.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL022327; CAB51559.1; -.
DR   INTERPRO; IPR001209; -.
DR   PROSITE; PS00527; RIBOSOMAL_S14; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   415 AA;  44978 MW;  C9D21BFF51CB2AC9 CRC64;


  Query Match              9.0%;  Score 85.5;  DB 4;  Length 415;
  Best Local Similarity   20.8%;  Pred. No. 2.5;
  Matches   32;  Conservative   31;  Mismatches   54;  Indels   37;  Gaps    6;

Qy      26 LWLFCSIVM--LLFLCSFSWLIFIFLQLETAKEPCMAKFGPLPSKWQMASSEPPCVNKVS 83
           :|  | :|:       :|: :: : |:|  |    :|           |||  |  |
Db     165 IWFGCKLVLNPSAININFNLILLLLLELLMAATVIIA----------ARSSEEDCKKK-- 212

Qy      84 DWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMIQTLTNKSKIQNVGGTYELH 143
                       | :: :||  |  ||  |: |:  ::: :   | :  :||   |:
Db     213 --------------KGSMSDSANILDEVPFPARVLKSYSVVEVIAGISAV--LGGIIALN 256

Qy     144 VGDTIDLIFNSEHQVLKNNTYWGIILLANPQFIS 177
           | |::    :  |  :   |:: |::   |  |:
Db     257 VDDSV----SGPHLSV---TFFWILVACFPSAIA 283


RESULT   5
Q65605
ID   Q65605      PRELIMINARY;      PRT;   579 AA.
AC   Q65605;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   SURFACE ENVELOPE PROTEIN (FRAGMENT).
OS   Bovine immunodeficiency-like virus.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11647;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OK;
RX   MEDLINE=96010247; PubMed=7571444;
RA   Suarez D.L., Whetstone C.A.;
RT   "Identification of hypervariable and conserved regions in the surface
RT   envelope gene in the bovine lentivirus.";
RL   Virology 212:728-733(1995).
DR   EMBL; L43128; AAC42158.1; -.
KW   Envelope protein.
```

```
FT   NON_TER       1      1
FT   NON_TER     579    579
SQ   SEQUENCE   579 AA;  64738 MW;  52EF25B35099A91D CRC64;


  Query Match             8.8%;  Score 84;  DB 12;  Length 579;
  Best Local Similarity   23.0%;  Pred. No. 5.1;
  Matches   47;  Conservative   28;  Mismatches   51;  Indels   78;  Gaps   11;

Qy      20 QRSSWKLWLFCSIVMLLFLCSFSWLIFI-----------------FLQL-ETAKEP--C 58
           :| : : |    ||:|| | |||| ::                  |  : :|:| |  |
Db     260 RRLTHREWSVSRIVVLL-LISFSWGMWTQRVEASGQHVAMVISPPGFRPVNDTSKIPWYC 318

Qy      59 MAKFGPLPS--------KWQMASSEPPCVNKVSDW----------KLEILQNGLYLIYGQ 100
           ::   |:|:         |:     |   |||:|||          : ::|:  :| :
Db     319 LSD-APIPTCRSPRWEQKYYQGKIEYEVVNKLSDWAEKHSRATWIEPDLLEEVVYEL--- 374

Qy     101 VAPNANYNDVAPFEVRLYKNKDMIQTLTNKSKIQNVGGTYELHVGDTIDLIFNSEHQVLK 160
                     |  |     |:|| : | |   | |   | |: |:       ||
Db     375 --------------ALLSANGSRQVTVTNGTDICNESDTPENQTGQTMTLLH------LK 414

Qy     161 N---NTY----------WGIILL 170
           |   ||:           |  :|:
Db     415 NQISNTWILNTSLTFCVQWPYVLI 438


RESULT   6
O96143
ID   O96143      PRELIMINARY;      PRT;  1188 AA.
AC   O96143;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PROTEIN WITH 5'-3' EXONUCLEASE DOMAIN (KEM-1 FAMILY).
GN   PFB0205C.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99021743; PubMed=9804551;
RA   Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA   Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA   Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA   Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA   Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT   "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Science 282:1126-1132(1998).
DR   EMBL; AE001380; AAC71830.1; -.
KW   Exonuclease.
SQ   SEQUENCE   1188 AA;  142895 MW;  BF767FC8532EBAC9 CRC64;


  Query Match             8.6%;  Score 82;  DB 5;  Length 1188;
  Best Local Similarity   30.4%;  Pred. No. 19;
  Matches   38;  Conservative   11;  Mismatches   36;  Indels   40;  Gaps    9;

Qy      41 FSWLIFIFLQLETAKEPCMAKFG-PLPSKWQMASSEPPCV-----NKVSDW----KLEIL 90
           : ||:|:|  ||       :|| | |   || : :: | ||     ||: ||    |||
Db       7 YMWLVFLFFFLE------LAKCGIPGLHKW-VINNFPSCVKIVDRNKLIDWNCIGKLEKA 59

Qy      91 Q---------------NGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMIQTLTNKSKIQN 135
           :               |    |     : ||||      |  ||||   |  |     |
Db      60 KGKHKRNHNGGDNNGDNNGDNNYDDNYDDNNYNDGCEIN-RNIKNKD--NTYDN-----N 111

Qy     136 VGGTY 140
           :  ||
Db     112 INNTY 116


RESULT   7
Q17701
ID   Q17701      PRELIMINARY;      PRT;  665 AA.
AC   Q17701;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   C06B3.2 PROTEIN.
GN   C06B3.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Percy C.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z77652; CAB01113.1; -.
SQ   SEQUENCE   665 AA;  75422 MW;  62EE7D5F14FB083F CRC64;


  Query Match             8.4%;  Score 79.5;  DB 5;  Length 665;
  Best Local Similarity   21.2%;  Pred. No. 16;
  Matches   44;  Conservative   24;  Mismatches   61;  Indels   79;  Gaps    9;

Qy      27 WLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMAKFGPLPSKWQMASSEPPCVNKVSDW- 85
           :|||   :| |:|| |  :| : ::|:       | |  |    :  | |   : |
Db       7 FLFC---LLKFVCSHSGYLFDYGKIES--------FEPFSSFINYSHVTPECTKDMKTWI 55

Qy      86 ----------------------KLEILQNGLYLI--YGQVA--PNANYNDVAPFEVRLYK 119
                                  | ||:: :| |  |   |  |:|   :|:     |:
Db      56 SSIEQFTAASSDCLFFKNCTASILNILKDNIYAIQQYDAFAKIPSAGLLEVSLVFDGSYQ 115

Qy     120 NKDMIQTL---TN--------------KSKIQNVGGTYELH----------VGDTIDL- 150
                |  :   ||                |||  | :                ||:||
Db     116 ECHRISGVKYETNYCYVLLVPGRNANCSGGSIQNAGTSVAFRRAVCMPKSCASSDTVDLY 175

Qy     151 ------------IFNSEHQVLKNNTYWG 166
                       | |:  | |:: :||
Db     176 NQLSALPLTACATFCSQRDVPKDSAFWG 203


RESULT   8
Q9LFL1
ID   Q9LFL1      PRELIMINARY;      PRT;  851 AA.
AC   Q9LFL1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   RECEPTOR PROTEIN KINASE-LIKE PROTEIN.
GN   F2K13_50.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
```

```
RA   Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL391141; CAC01703.1; -.
KW   Kinase.
SQ   SEQUENCE   851 AA;  95721 MW;  97CF0A4B4188EB64 CRC64;


  Query Match             8.4%;  Score 79.5;  DB 10;  Length 851;
  Best Local Similarity   22.2%;  Pred. No. 22;
  Matches   41;  Conservative   25;  Mismatches   44;  Indels   75;  Gaps    8;

Qy      43 WLIFIFLQLETAKEPCMAKFGPLPSKWQMASSEPPCVNKV------SD------------ 84
           :| |||  :   :    |: | :     : |:||| :  |      ||
Db       7 YLFFIFAIIHYVQ----AQQGFISLDCGLPSNEPPYIEPVTGLVFSSDADHIPSGISGRI 62

Qy      85 -------------------------WKLEILQNGLYLI-----YGQVAPNANYNDVAPFE 114
                                     : |::||| |:|     ||       |||   |:
Db      63 QKNLEAVHIKPYLFLRYFPDGLRNCYTLDVLQNRRYMIKAVFVYGNY---DGYNDYPSFD 119

Qy     115 VRLYKNKDMIQTLTNKSKIQNVGGTYE--LHV-------------GDTIDLIFNSEHQVL 159
           : |  || :   |  |     | |: |  :|:             |::  |   | ::|
Db     120 LYLGPNKWVRVDLEGK-----VNGSVEEIIHIPSSNSLQICLVKTGNSLPFISALELRLL 174

Qy     160 KNNTY 164
           :|:||
Db     175 RNDTY 179


RESULT   9
P89064
ID   P89064      PRELIMINARY;      PRT;   326 AA.
AC   P89064;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   STRAIN 1040 GENE 9.
GN   9.
OS   Rotavirus.
OC   Viruses; dsRNA viruses; Reoviridae.
OX   NCBI_TaxID=10912;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1040;
RA   Das M.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X95372; CAA64659.1; -.
DR   INTERPRO; IPR001963; -.
DR   PFAM; PF00434; VP7; 1.
SQ   SEQUENCE   326 AA;  37287 MW;  86589AED9B457076 CRC64;


  Query Match             8.2%;  Score 78;  DB 12;  Length 326;
  Best Local Similarity   22.3%;  Pred. No. 9.8;
  Matches   41;  Conservative   30;  Mismatches   59;  Indels   54;  Gaps    9;

Qy      26 LWLFCSIVMLLFLC-----SFSWLIFIFLQLETAKEPCMAK---------FGPLPSKWQM 71
           | :  ||::| ::       :  ::|| || |     | :            | | : :
Db      10 LTILISIILLNYILKTITNTMDYIIFRFLLLIALMSPFVRTQNYGMYLPITGSLDAVYTN 69

Qy      72 ASSEPP------CV-------NKVSDWKLEILQNGLYLIYGQVAPNA---NYNDVAPFEV 115
           ::||        |:       |::|| : |   : |:|  |    :    :|||:  | :
Db      70 STSEESFLTSTLCLYYPTEAKNEISDNEWENTLSQLFLTKGWPTGSVYFKDYNDITTFSM 129

Qy     116 --RLYKNKDMIQTLTNKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANP 173
             :|| : :::        :     | ||   :  ||| |           |    | ||
Db     130 NPQLYCDYNVV--------LMRYDNTSELDASELADLILNE----------W----LCNP 167

Qy     174 QFIS 177
             ||
Db     168 MDIS 171


RESULT  10
P75041
ID   P75041      PRELIMINARY;      PRT;   564 AA.
AC   P75041;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HEXOSEPHOSPHATE TRANSPORT PROTEIN.
GN   UHPT.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
DR   EMBL; AE000010; AAB95727.1; -.
SQ   SEQUENCE   564 AA;  61631 MW;  EE65E57D11C8F953 CRC64;


  Query Match             8.2%;  Score 78;  DB 2;  Length 564;
  Best Local Similarity   22.9%;  Pred. No. 19;
  Matches   33;  Conservative   22;  Mismatches   45;  Indels   44;  Gaps    8;

Qy      17 QGAQRSSWKLWLFCSIVMLL----FLCSFSWLIFIFLQLETAKEPCMAKFGPLPSKWQMA 72
           |  |::  | |:|    ::|     || |:|       |   :| | :     :| ||
Db     238 QSVQQAGTKYWVFILTALILIGFGILCVFAW-------FEKQMDPFMPQ---KQTKEQMQ 287

Qy      73 SSEPPCVNKV----SDWK--------LEILQN----GLYLIYGQVAPNANYNDVAPFEVR 116
                |    :   : ||         | :| |    | : |   |:| :::|
Db     288 LGNQPSAGDILKRKATWKMIGMYGICLVVLVNPLTGGWWNILQAVSPASSFN-------- 339

Qy     117 LYKNKDMIQTLTNKSKIQNVGGTY 140
               || ::||     ::  || :
Db     340 ---VKDGVKTL---KPLEGAGGYF 357


RESULT  11
Q9KM00
ID   Q9KM00      PRELIMINARY;      PRT;   352 AA.
AC   Q9KM00;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PEPTIDE ABC TRANSPORTER, PERMEASE PROTEIN, PUTATIVE.
GN   VCA0589.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleishmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004389; AAF96491.1; -.
```

```
DR   TIGR; VCA0589; -.
SQ   SEQUENCE   352 AA;  39588 MW;  CA9A777BE480D3B9 CRC64;


  Query Match             8.1%;  Score 77.5;  DB 2;  Length 352;
  Best Local Similarity   26.3%;  Pred. No. 12;
  Matches   31;  Conservative   19;  Mismatches   35;  Indels   33;  Gaps    8;

Qy      10 PLSH---SRTQGAQRSSWKLWLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMAKFGPLP 66
           ||:    :| :  :|  | ||:|    :|||: |      :| :|    :| : ::
Db      14 PLTEARWARFKANRRGFWSLWIF----LLLFVVS------LFAELIANDKPLLIQY---D 60

Qy      67 SKWQMASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMI 124
             | |    | | : |:      | |     |:    |:| |  |: | | : |  |
Db      61 GAWYM-----PIVQRYSE-----TQFG----GEFDTEADYTD--PYVVSLIEEKGQI 101


RESULT  12
Q9W4E2
ID   Q9W4E2      PRELIMINARY;      PRT;  3614 AA.
AC   Q9W4E2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   AKAP550 PROTEIN.
GN   AKAP550.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003433; AAF46011.1; -.
DR   FLYBASE; FBgn0021748; Akap550.
DR   INTERPRO; IPR000104; -.
DR   INTERPRO; IPR000179; -.
DR   INTERPRO; IPR000409; -.
DR   INTERPRO; IPR001680; -.
DR   PFAM; PF00400; WD40; 3.
DR   PFAM; PF02138; Beach; 1.
DR   PRINTS; PR00308; ANTIFREEZEI.
DR   PROSITE; PS00193; CYTOCHROME_B_QO; UNKNOWN_1.
SQ   SEQUENCE   3614 AA;  400162 MW;  621563EF59703177 CRC64;


  Query Match             8.1%;  Score 77.5;  DB 5;  Length 3614;
  Best Local Similarity   21.7%;  Pred. No. 1.9e+02;
  Matches   33;  Conservative   26;  Mismatches   62;  Indels   31;  Gaps    5;

Qy      27 WLFCSIVMLLFLCSFSWLIFIFLQLETAKEPCMA--------KFGPLPSKWQMASSEPPC 78
           :|  :: : :|| | : ::| |     | |:   |         |:| :|   : :   |
Db    2887 YLLQNVALEIFLASRTSILFAFPDQHTVKKVIKALPRVGVGIKYG-IPQTRRASMMSPRQ 2945

Qy      79 V----NKVSDWKLEILQNGLYLIYGQVAPNANYND-----VAPFEVRLYKNKDMIQTL-- 127
           :    |    |:   : |  ||::         |||     : |: :  |::||:  :|
Db    2946 LMRNSNMTQKWQRREISNFEYLMFLNTIAGRTYNDLNQYPIFPWVLTNYESKDLDLSLPS 3005

Qy     128 ----------TNKSKIQNVGGTYELHVGDTI 148
                     | |:      ||    |||
Db    3006 NYRDLSKPIGALNPSRRAYFEERYESWDSDTI 3037


RESULT  13
Q9QS10
ID   Q9QS10      PRELIMINARY;      PRT;   326 AA.
AC   Q9QS10;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VP7 PROTEIN.
GN   VP7.
OS   Human rotavirus serotype G2.
OC   Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX   NCBI_TaxID=73034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TA3;
RX   MEDLINE=99301494; PubMed=10374958;
RA   Zao C.L., Yu W.N., Kao C.L., Taniguchi K., Lee C.Y., Lee C.N.;
RT   "Sequence analysis of VP1 and VP7 genes suggests occurrence of a
RT   reassortant of G2 rotavirus responsible for an epidemic of
RT   gastroenteritis.";
RL   J. Gen. Virol. 80:0-0(0).
DR   EMBL; AF106280; AAD47299.1; -.
DR   INTERPRO; IPR001963; -.
DR   PFAM; PF00434; VP7; 1.
SQ   SEQUENCE   326 AA;  37174 MW;  4782203A59C6A8A5 CRC64;


  Query Match             8.1%;  Score 77;  DB 12;  Length 326;
  Best Local Similarity   22.8%;  Pred. No. 12;
  Matches   42;  Conservative   29;  Mismatches   59;  Indels   54;  Gaps    9;

Qy      26 LWLFCSIVMLLFLC-----SFSWLIFIFLQLETAKEPCMAK---------FGPLPSKWQM 71
           | :  ||::| ::       :  ::|| || |      | :.           | | : :
Db      10 LTILISIILLNYILKTITNTMDYIIFRFLLLIALISPFVRTQNYGMYLPITGSLDAVYTN 69

Qy      72 ASSEPP------CV-------NKVSDWKLEILQNGLYLIYGQVAPNA---NYNDVAPFEV 115
           ::|  |       |:       |::|| : |    : |:|  |     :    :|||:  | |
Db      70 STSGEPFLTSTLCLYYPAEAKNEISDDEWENTLSQLFLTKGWPIGSVYFKDYNDINTFSV 129

Qy     116 --RLYKNKDMIQTLTNKSKIQNVGGTYELHVGDTIDLIFNSEHQVLKNNTYWGIILLANP 173
             :|| : :::        :     | ||   :  ||| |           |    | ||
Db     130 NPQLYCDYNVV--------LMRYDNTSELDASELADLILNE----------W----LCNP 167
```

```
Qy     174 QFIS 177
           ||
Db     168 MDIS 171
```

```
RESULT  14
Q26197
ID   Q26197      PRELIMINARY;      PRT;   433 AA.
AC   Q26197;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   SERA-1 (FRAGMENT).
OS   Plasmodium vinckei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5860;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rosenthal P.J., Gor D.O., Li A.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U59860; AAC61622.1; -.
DR   MEROPS; C01.077; -.
DR   INTERPRO; IPR000668; -.
DR   PRINTS; PR00705; PAPAIN.
FT   NON_TER       1      1
SQ   SEQUENCE   433 AA;  50159 MW;  743287DBB22B3279 CRC64;
```

```
  Query Match             8.1%;  Score 77;  DB 5;  Length 433;
  Best Local Similarity   21.3%;  Pred. No. 17;
  Matches   42;  Conservative   20;  Mismatches   59;  Indels   76;  Gaps   10;

Qy      39 CSFSWLIFIFLQLETAKEPCMAKFG---------------------------------- 63
           ||  ||    | |||  :  ||  ||
Db      30 CSLCWLFASKLHLETIR--CMRGFGHNRGSALYVANCSERTADNVCDVGSNPIEFLKILE 87

Qy      64 -----PL----PSKWQMASSEPPCVNKVSDW-----KLEILQNGLYLIY----GQVAPNA 105
                ||    |  |:  |   |    ::|       ::| | :|  :    | :| |:
Db      88 KNKFLPLESNFPYVWKNVSD--TCPKPANNWTNIWGNTKLLYNNMYGEFIQHRGFIAYNS 145

Qy     106 NY--------NDVAPFEVR-------LYKNKDMIQTLTNKSKIQNVGGTYELHVGDTIDL 150
           |         |:   |:|         | :|:|:   |   | |: |    |    :::
Db     146 RYFAKNMDVFIDIVKREIRNKGSVIAYIKTRDVIEYDFNGRYISNICG--HSHPDHAVNI 203

Qy     151 IFNSEHQVLK--NNTYW 165
           |    :   |    |||
Db     204 IGYGNYISDKGEKRTYW 220
```

```
RESULT  15
Q9P3D8
ID   Q9P3D8      PRELIMINARY;      PRT;   748 AA.
AC   Q9P3D8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROBABLE DOLICHYL-PHOSPHATE-MANNOSE--PROTEIN MANNOSYLTRANSFERASE.
GN   B13I18.50.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL390189; CAB99175.1; -.
SQ   SEQUENCE   748 AA;  85290 MW;  A91DB796F41CB9C9 CRC64;
```

```
  Query Match             8.1%;  Score 77;  DB 3;  Length 748;
  Best Local Similarity   21.2%;  Pred. No. 33;
  Matches   38;  Conservative   25;  Mismatches   66;  Indels   50;  Gaps    8;

Qy      13 HSRTQGAQRSSWKLWLFCSIVMLLFLCSFSWL-IFIFLQLETAKEPCMAKFGPLPSKWQM 71
           |:  : :    |  ||| : : : |:||  |: :|:          :    :   |
Db     222 HNLRKKSFTPEWAFWLFLTGLSIGFVCSVKWVGLFV---------TALVGLYTIEDLW-- 270

Qy      72 ASSEPPCVNKVSDWKLEILQNGLYLIYGQVAPNANYNDVAPFEVRLYKNKDMIQTLTNK- 130
                 ||  | |: :      ::   ||   :   : || | |         |||
Db     271 --------NKFGDTKMPV-----KVLAAHVAARVSCLIILPFAVYLLSFAIHFWVLTNSG 317

Qy     131 -------SKIQ-NVGGT-------YELHVGDTIDLIFNSEHQVLKNNTYWGIILLANPQ 174
                 |  | |: ||        || :|  :          :||  | | :| :: |
Db     318 PGDAQMPSLFQANLRGTEVGRDSPLELAIGSRV---------TIKNMGYGGGLLHSHVQ 367
```

Search completed: February 23, 2001, 10:49:18
Job time: 33 sec